08/187,879

Europäisches Patentamt
European Patent Office
Office européen des brevets

(19) (11) Publication number: **0 293 792 A2**

(12) **EUROPEAN PATENT APPLICATION**

(21) Application number: 88108569.0

(22) Date of filing: 28.05.88

(51) Int. Cl.⁴: **C12N 15/00, C12N 1/20, A61K 39/21, G01N 33/569**

The applicant has filed a statement in accordance with Rule 28 (4) EPC (issue of a sample only to an expert). Accession number(s) of the deposit(s): ATCC 67378.

(30) Priority: 01.06.87 US 57569

(43) Date of publication of application: 07.12.88 Bulletin 88/49

(84) Designated Contracting States: AT BE CH DE ES FR GB GR IT LI LU NL SE

(71) Applicant: E.I. DU PONT DE NEMOURS AND COMPANY
1007 Market Street
Wilmington Delaware 19898(US)

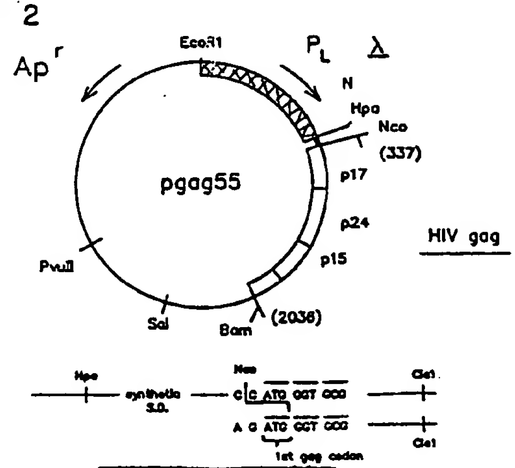
(72) Inventor: Ferguson, Blair Quinn
2513 W. 18th Street
Wilmington Delaware 19806(US)
Inventor: Tribe, David Edward
431 Dean Drive
Kennett Square Pennsylvania 19348(US)
Inventor: Petteway, Stephen Robert
1017 Ashley Street
West Chester Pennsylvania 1382(US)

(74) Representative: von Kreisler, Alek, Dipl.-Chem. et al
Patentanwälte Von Kreisler-Selting-Werner
Deichmannhaus am Hauptbahnhof
D-5000 Köln 1(DE)

(54) **Human immunodeficiency virus GAG-encoded proteins.**

(57) The invention concerns nonfused recombinant proteins corresponding to the complete HIV gag and the complete HIV p17gag subregion coding sequences, which are useful immunological components of diagnostics, therapeutics, and vaccines.

F I G. 2



EP 0 293 792 A2

Human Immunodeficiency Virus gag-encoded Proteins

FIELD OF THE INVENTION

The present invention concerns proteins derived by molecular cloning that are useful immunological components of diagnostics, therapeutics, and vaccines for infectious diseases. In particular, the invention relates to proteins encoded by the gag gene of human immunodeficiency virus, the etiologic agent of acquired immunodeficiency syndrome.

BACKGROUND

The human immunodeficiency virus (HIV, also LAV, HTLV-III, or ARV) is the primary etiologic agent of the acquired immune deficiency syndrome (AIDS) [Barre-Sinoussi et al., Science 220:868-871 (1983); Gallo et al., Science 224:500-503 (1984); Levy et al., Science 225:840-842 (1984)]. The underlying disease state involves a tropism of HIV for the $T4^{+}$ lymphocyte subset resulting in a selective depletion of the helper/inducer cells of the immune system, leaving the HIV-infected individual defenseless against a number of opportunistic infections. Approximately two million people in the United States and five million or more individuals world-wide may now be infected by HIV. The U.S. Public Health Service now estimates that the cost of caring for AIDS patients in 1991 will be between$8 billion to $16 billion per year. Thus, the development of diagnostics, therapeutics, and vaccines to HIV is the subject of intense medical research.

The nucleotide sequence of several independent viral isolates of HIV have been determined [Ratner et al., Nature 313:277-284 (1985); Muesing et al., Nature 313:250-258 (1985); Wain-Hobson et al., Cell 40:9-17 (1985); Sanchez-Pescador et al., Science 227:484-492 (1985)]. HIV is related to the lentivirus subgroup of retroviruses. Lentiviruses are associated with slowly developing, non-tumorigenic diseases in other animal systems. The major structural proteins of the HIV infectious particle are encoded by the viral gag, pol and env genes. In addition to the gag,pol, and env genes, the genome of HIV contains several other open reading frames designated sor,tat, art/trs and 3'-orf which encode additional viral proteins. These proteins are known to serve important regulatory functions during the HIV infectious cycle.

The gag gene products are involved in the structure of the genomic RNA-containing core and the assembly of the virus. The gag gene is termed "gag" for group-specific antigen, since in other retroviral systems the analogous gene encodes proteins that are cross- reactive with similar proteins of a related group of retroviruses. The HIV gag gene encodes a precursor of about 55,000 daltons (55 kDa) which is designated $p55^{gag}$. The gag coding sequence of HIV isolate BH10 (HIV_{BH10}) (Ratner et al., op. cit.) is contained within nucleotides 334 to 1869 and is comprised of 512 codons (Figure 1). The nucleotide numbering is according to Ratner et al. (op. cit.).

During the HIV infectious life cycle, the gag precursor $p55^{gag}$ is proteolytically processed into species of approximately 24, 17, and 15 kDa. The gag proteins are designated $p24^{gag}$, $p17^{gag}$, and $p15^{gag}$. The mature gag proteins are derived from subregions of the $p55^{gag}$ precursor in the order N-p17-p24-p15-C (N- and -C designate the amino and carboxyl terminus of the protein, respectively).

As is the case for the N-terminal gag-encoded proteins of other mammalian retroviruses [Henderson et al., Proc. Natl. Acad. Sci. 80:339-343 (1983); Rein et al. Proc. Natl. Acad. Sci. 83:7246-7250 (1986)], the N-terminal gag-encoded $p17^{gag}$ is modified at its mature N-terminal glycine by myristylation. The N-terminal lipid myristic acid moiety of $p17^{gag}$ presumably mediates the interaction of $p17^{gag}$ with the lipid membrane of the virus particle. Gag-derived determinants are known to be exposed on the surface of cells infected by other mammalian retroviruses [Schiff-Maker and Rosenberg, Virology 154:286-301 (1986)]. Thus, HIV $p17^{gag}$-specific epitopes may be present on the surface of infectious HIV particles and HIV-infected cells. The $p24^{gag}$ protein is the major virus capsid or core protein. The C-terminal gag-encoded $p15^{gag}$ protein is highly basic and likely constitutes the core ribonucleoprotein and binds nonspecifically to many sites on the viral RNA.

Earlier reports describe the expression of subregions of HIV gag protein in heterologous expresion systems such as in E. coli [Dowbenko et al., Proc. Natl. Acad. Sci. 82: 7748-7752 (1985); Chang et al., Science 228:93-99 (1985); Ghrayeb et al., DNA 5:93-99 (1986); Steimer et al., Virology 150: 283-290 (1986); Shoeman et al., Anal. Biochem. 161:370-379 (1987)] and in yeast [Kramer et al., Science 231:1580-1584 (1986)]. These recombinant gag-derived proteins produced in heterologous expression systems were shown to be useful as antigens for the detection of HIV gag-specific antibodies in human sera. No previous report has described the production using a het-

erologous expression system, for example an *E. coli* plasmid expression system, of proteins corresponding to the full length, nonfused *gag* precursor $p55^{gag}$ or to the mature *gag* subregion $p17^{gag}$. The present invention provides the full length nonfused HIV *gag* precursor $p55^{gag}$ and the mature *gag* subregion $p17^{gag}$ as produced using a heterologous expression system.

SUMMARY O[illegible] INVENTION

The invention provides proteins corresponding to the full length, nonfused HIV *gag* precursor $p55^{gag}$ and the mature nonfused $p17^{gag}$ subregion of $p55^{gag}$ which are produced using molecular cloning methods and heterologous expression vector systems. Polypeptides [illegible]responding to $p55^{gag}$ and $p17^{gag}$ of the HIV isolate BH10 (Ratner et al., *op. cit.*) were expressed in *E. coli*. Protein $p55^{gag}$ is encoded by nucleotides number 334 to 1869 (nucleotide numbering according to Ratner et al., *op. cit.*) (see Figure [illegible]). The $p17^{gag}$ segment of $p55^{gag}$ is encoded by nucleotides number 334 to 729. The HIV *gag* proteins are useful immunological components of [illegible] to identify individuals and blood [illegible] that have been exposed to HIV, of reagents [illegible] can be used to monitor and stage the [illegible] of the disease in clinical therapeutic trials [illegible] agents able to control the disease [illegible] vaccine that will protect iduals who are exposed to HIV.

DETAILED DESCRIP[illegible] THE INVENTION

Two recombinant [illegible] containing only amino acid sequences [illegible] g to the *gag* coding sequence of HIV [illegible] created. These recombinant *gag*-derived polypeptides are highly immunoreactive with [illegible] sera from individuals infected with [illegible] term "peptide" is well known in the art [illegible] to a compound of two or more amino acids. One of the peptides of the invention correspo[illegible] the complete and nonfused HIV *gag* [illegible]cursor protein. In HIV_{BH10}, $p55^{gag}$ is en[illegible] nucleotides 334 to 1869 (nucleotide num[illegible] according to Ratner et al., *op. cit.*). A secon[illegible] of the invention corresponds to the [illegible]-terminal *gag*-encoded protein, $p17^{ga}$[illegible] $p17^{gag}$ is encoded by nucleotides [illegible] to 7[illegible] (Ratner et al., *op. cit.*). The specific domai[illegible] V are known in the art.

The present *E.* [illegible] sed proteins corresponding to HIV p5[illegible] 17^{gag} are referred to as GAG55 and GAG17, respectively. The structures of plasmid expression vectors designed for the expression of GAG55 and GAG17 in *E. coli* are detailed in Example 1.

As used herein, the term "expression vector" includes a DNA molecule which contains signals, recognized by a particular host biological cell, that direct and determine the expression of a particular gene sequence in the host cell. The vector DNA may, for example, contain information that determines the uptake of the vector DNA by the host cell, the integration and replication of the vector DNA, and the corresponding RNA biosynthesis and translation. Examples of expression vectors include recombinant viruses, plasmids and genes.

It is to be understood that the term "corresponding to" includes modifications of the specified amino acid sequences which do not adversely affect the antigenic characteristics of the peptide of the invention. For example, different isolates of HIV are known to differ to some extent in the predicted amino acid sequence of the *gag*-derived proteins (Ratner et al., *op. cit.*; Muesing et al., *op. cit.*; Wain-Hobson et al., *op. cit.*; Sanchez-Pescador, *op. cit.*). One skilled in the art could align the amino acid sequences of peptides from different sources to the schematic of Figure 1 or the nucleotide sequence of Ratner et al. (*op. cit.*) to identify the segments therein which correspond to the peptides defined herein.

HIV *gag*-derived proteins are useful immunological components of diagnostics, therapeutics, and vaccines for HIV-related disease, including acquired immune deficiency syndrome (AIDS). HIV-expressed *gag* proteins are known to be useful immunological antigens for the detection of HIV *gag*-specific antibodies [Petricciani, *Ann. Int. Med.* 103:726-729 (1985)]. The analysis of HIV-specific antibodies in human serum samples is extremely important in diagnosis of disease and for screening of human blood products. An enzyme-linked immunosorbent assay (ELISA) which uses HIV grown in tissue culture cells as antigen is now widely used for detecting sera positively immunoreactive to HIV (Petricciani, *op. cit.*). The ELISA derived from purified HIV particles grown in tissue culture is designated HIV-ELISA.

Since HIV infection leads to a series of clinical manifestations accompanied by variable antibody levels to different antigens, useful blood screening and diagnostic reagents should contain multiple antigens. Moreover, it is desirable to detect and quantitate antibody levels to specific antigenic components of HIV. Although the HIV-ELISA is a sensitive tool for the detection of HIV-specific antibodies in human serum samples, this test does not provide information as to which antigenic components of HIV are reactive with human antibodies.

The production of individu[illegible] defi[illegible] HIV-derived proteins, such as gag-enco[illegible] p[illegible]ins, in E. coli provides antigen reagents [illegible] the [illegible]elective detection and quantitation of anti[illegible]ies [illegible]cognizing specific antigens, such as gag-[illegible]od[illegible] antigens. HIV gag proteins derived by [illegible] [illegible]ecu[illegible] cloning and expressed in heterologous ex[illegible]ress[illegible] systems, i.e., recombinant gag proteins, a[illegible] for the development of novel HIV diagn[illegible] For example, the E. coli produced HIV G[illegible]55 [illegible] GAG17 proteins can be efficiently u[illegible]ed [illegible] a diagnostic ELISA for the detection a[illegible]qua[illegible]titation of HIV gag-specific antibodies that [illegible] as[illegible]iated with prior infection with HIV.

A heterologous expressi[illegible] sys[illegible] such as that utilizing an E. coli plasmid [illegible] vector, provides a preferred alterna[illegible] [illegible]fectious HIV grown in tissue culture as [illegible] ur[illegible] of antigen for ELISA blood screening a[illegible] di[illegible]stic reagents. Importantly, the E. coli plas[illegible] [illegible]ession system is free of infectious HIV. M[illegible]ver, [illegible] gag-derived proteins GAG55 and GAG[illegible] [illegible]an [illegible] expressed at high levels in E. coli using [illegible] [illegible]riate vectors, thereby facilitating the purif[illegible] these useful polypeptides. Thus, recomb[illegible] can be produced in large quantities [illegible] [illegible]he biohazard and containment problems [illegible] with the production of gag proteins fro[illegible] HIV virions.

The recombinant gag [illegible] [illegible]f the present invention are preferred to [illegible] described recombinant gag proteins. The [illegible]inant product GAG55 corresponds to the [illegible]th, non-fused product of the complete [illegible] [illegible]g sequence. Previously described re[illegible] [illegible]ag proteins only contain subregions o[illegible] [illegible]xample, in a recombinant gag protein [illegible] by Kramer et al. (op. cit.) and Shoema[illegible] [illegible] cit.), amino acids number 2 to 12 of [illegible] [illegible]eleted. Other reported recombinant gag [illegible] reported have larger segments of the [illegible] [illegible]ence deleted (Dowbenko et al., op. cit.; [illegible] al., op. cit.; Ghrayeb et al., op. cit.; St[illegible] op.cit.). More gag-specific antigenic de[illegible] [illegible]re represented by the product of the [illegible] [illegible]ll length gag coding sequence than a[illegible] [illegible]ed by incomplete subregions of gag. [illegible] [illegible]roduct of the full length gag coding se[illegible] [illegible]h as GAG55, is expected to be a more [illegible] [illegible]agent for the detection of gag-specific [illegible] [illegible]an are antigens comprised of only s[illegible] gag.

In particular, GAG55 [illegible] [illegible]7 contain the region of gag corresp[illegible] [illegible]he complete $p17^{gag}$. This complete regi[illegible] [illegible]s not present in other reported recom[illegible] proteins. As discussed previously an[illegible] [illegible]$p17^{gag}$ region of gag contains antigenic [illegible] that may be presented on the surfac[illegible] HIV-infected cells, and may be crit[illegible] [illegible]otective host immune response to HIV.

Another advantage of the peptides of the invention is that the recombinant gag-derived proteins are not fused to other, non-HIV, or non-HIV gag sequences. The inclusion of additional non-HIV sequences, such as E. coli-derived sequences, fused to the HIV gag coding sequence is not desirable since these sequences could lead to frequent false-positive (i.e., non-HIV) antibody reactivity. Thus, the HIV gag sequences have been expressed in E. coli in the absence of fused bacterial or other non-HIV protein sequences. It is also desirable to express nonfused authentic gag proteins, rather than fusion proteins, in order to facilitate the correct folding and native structure of the product. The correctly folded structure of the product may be critical to its effectiveness as a diagnostic, therapeutic, or vaccine reagent. Previously reported recombinant HIV gag-derived proteins are fusion polypeptides containing non-HIV sequences.

Although most current strategies for the development of an AIDS vaccine are focusing on the HIV env gene, the gag antigens may well be an important component of a protective vaccine against HIV. Indeed, recent reports suggest that gag $p17^{gag}$ antigenic determinants are accessible to the neutralizing effects of specific antibodies [Sarin et al., Science 232: 1135- 1137 (1986)]. As discussed previously, $p17^{gag}$ determinants may be exposed on the envelope of the virus itself, since the gag $p17^{gag}$ protein is thought to be associated with the virus lipid membrane.

AIDS patients are less likely to have antibodies to gag-encoded proteins compared to individuals with milder forms of the disease [Steimer et al., Virology 150: 283-290 (1986)]. Thus, there is at least correlative evidence that the gag-specific host immune response may be critical for the effective control of HIV by an HIV-exposed individual.

Due to the heterogeneity in the sequence of env from various HIV isolates [Ratner et al., Nature 313:277-284 (1985); Wain-Hobson et al., Cell 40:9-17 (1985); Muesing et al., Nature 313:250-258 (1985); Sanchez-Pescador et al., Science 227:484-492 (1985)], it is possible that a vaccine containing env from a single HIV isolate may not have a broadly neutralizing effect on other HIV species. The degree of heterogeneity in the gag gene product among different HIV isolates is far less than that found for env [Ratner et al., Nature 313:277-284 (1985); Muesing et al., Nature 313:250-258 (1985); Wain-Hobson et al., Cell 40:9-17 (1985); Sanchez-Pescador et al., Science 227:484-492 (1985)]. Thus, HIV gag gene products may be an important component of a vaccine that is broadly protective against many isolates of HIV.

Since the peptides of the invention are produced by recombinant DNA technology in the ab-

sence of infectious H /. ... y provide several advantages over more ... ditior l vaccine approaches using killed or atte... ate... preparations. Recombinant peptides ... repare and administer and no genetic infor... is introduced by vaccination. In additi... r... inant peptides are more readily produce... and ... ...fied than the viral-expressed peptides ...solat... from virus-infected cells.

In one embodim... , th... peptide of the invention is used in a di... ... used for detecting antibodies to HIV in ... bi... sample. The peptide can be employ... in ... ...ess for detecting antibodies to HIV c...pri... ...ntacting a biological sample with the pept... ...nd detecting immunoreactivity. The p... ... can also be employed as a component in ... v... ...ine protective against HIV. The vaccine co... ri... ...n effectively protective amount of the p... ...

E. coli strain KA... 29... ...rmed with plasmid pGAG55, encoding ... ...tei... ... 55 and described in Example 1, has ... ...d with the American Type Culture ... olle... ... ATCC), Rockville, Maryland, and bears d... ...t accession number 67378. This deposit is av... ...le t... the public upon the grant of a patent ... ...ssi...nee. However, it should be understoo... ... ...vailability of a deposit does not con... ... ...e to practice the subject matter in de... ... ...tent rights granted by governmental ... tio...

Materials and Method...

Unless otherwise ... ... arts and percentages are by weight a... ... ...re Celsius.

Plasmid Constructio...

Plasmid constru... ... ... arried out using standard methodolog... ... ...crib...d by Maniatis et al., Molecular Clonin... ... ...y Plasmid, Cold Spring Harbor Lab... ... York (1982), the teaching of which is ... porated by reference. Enzymes and ... ...ts used for plasmid constructions ... ... from Bethesda Research Laborator... , ...urg, MD or New England Biolabs, Be... ... ...thods for digesting, identifying, reco... ... ...urifying the various nucleotide seq... ... in the invention are known to those ... ...rt as are methods for ligating the ... ...o vectors, transforming host micro ... ...s, cloning, and recovering product... ... Accordingly, the methods will only be d... by reference to specific embodimer... ...ention set forth hereinafter.

E. coli Strains, Plasmid Expression Vectors, and Induction of Expression

Several types of E. coli plasmid expression vectors were used for the production of HIV gag-derived proteins in E. coli. In the vectors used, expression of the desired product is controlled transcriptionally using either the operator and promoter of the E. coli tryptophan operon [Ivanoff et al., Proc. Natl. Acad. Sci. 83:5392-5396 (1986)] or the phage lambda P_L promoter [Rosenberg et al., Methods Enzymol. 101:123-138 (1983)]. We have also used the E. coli lactose operon promoter to direct expression of HIV gag-derived proteins in E. coli.

Plasmid vectors utilizing the phage lambda promoter were derived from plasmid pBF106. The E. coli plasmid expression vector pBF106 was constructed as follows. The small HpaI to BamHI fragment of plasmid pKC30 [Rosenberg et al., Methods Enzymol. 101:123-138 (1983); Rao, Gene 31:247-150 (1984)] (obtained from K. Abremski, Du Pont Experimental Station, Wilmington, DE) was replaced with the following oligonucleotide linker:

5′- AAC GAA TCC GAA GTG TAA GCC ATG -3′
3′- TTG CTT AGG CTT CAC ATT CGG TAC CTAG -5′

The inserted sequence provides signals for efficient translation initiation and a unique NcoI restriction endonucleose site. Since the translation start ATG forms part of the NcoI site, coding sequences can easily be fused in-frame directly to the ATG and translation initiation signal.

Expression from the P_L-containing plasmids is controlled transcriptionally using a lysogenic host that provides lambda cI repressor. Expression is induced by inactivating cI using nalidixic acid [Mott et al., Proc. Natl. Acad. Sci. 82:88-92 (1985)] or by temperature-shift and using a temperature sensitive repressor such as cI857 [Rosenberg et al., Methods Enzymol. 101:123-138 (1983)]. Temperature-shift inductions were carried out using host strain KA1298 (obtained from K. Abremski, Du Pont Experimental Station, Wilmington, DE), according to the procedure of Young et al., Proc. Natl. Acad. Sci., 80:6105-6109 (1983). KA1298 is a defective lambda lysogenic containing cI857. Nalidixic acid inductions were carried out using the defective lysogenic host DC550 (obtained from D. Court, NCI, Frederick, MD), according to Mott et al., Proc. Natl. Acad. Sci. 82:88-92 (1985).

E. coli plasmid expression vectors containing the tryptophan promoter were derived from pKGP36.trp (Ivanoff et al., op. cit). E. coli host strains MM294 and HB101, DH5 (F^- recA1, endA1, gyrA96, hsdR17, supE44), DH5Δ (lac), SC122, CAG456(htpR165), CAG629(lon-htpR165), and SG4119 (Δlon) were used [Maniatis et al., Molecu-

lar Cloning: A Laboratory Ma[illegible]al, Cold Spring Harbor Laboratory, NY (1982); [illegible]er et [illegible] Proc. Natl. Acad. Sci. 81:6779-6783]. [illegible] [illegible]m the trypophan promoter was carri[illegible]d o[illegible] described (Ivanoff et al., op. cit.).

Serum Samples and Serology

Human sera were coll[illegible] [illegible]i at [illegible] [illegible]od Bank of Delaware blood donor cen[illegible]s [illegible] [illegible]86 or were obtained from T. Mathews Duke [illegible] [illegible]ity. Serum was stored at -70°C befc[illegible]e testi[illegible] [illegible]d repeated freezing and thawing was avoid[illegible] [illegible] preliminary classification of all sera was ma[illegible] [illegible]ased on repeated tests with an HIV-ELI[illegible]A (Du P[illegible]nt, Wilmington, DE) blood screening ki[illegible], on the results of HIV immunoblot testing carrie[illegible] [illegible]t by [illegible]commercial laboratory (Biotech Resea[illegible] [illegible] [illegible]ies, Rockville, MD), and on testing [illegible] [illegible] [illegible]vity to env using a recombinant antig[illegible]n (E[illegible] [illegible]sed ELISA (described below). Proce[illegible]ure [illegible] [illegible]3-1/2 hour assay described in the Du Pon[illegible] [illegible] - III ELISA blood screening kit, was us[illegible] to [illegible] [illegible]era. Further immunoblot analysis was ca[illegible]ried out using commercially available HIV ant[illegible] [illegible]-cont[illegible]ng nitrocellulose strips (Du Pont-Bio[illegible] [illegible] Re[illegible] [illegible] Laboratories).

Immunoblot Analysis

Standard methods for p[illegible] [illegible]acrylam[illegible]de gel electrophoresis (SDS-PAGE) [L[illegible] [illegible]mli, N[illegible]ture 227:680-685 (1970)] and electroph[illegible] [illegible] of proteins to nitrocellulose [Towbin [illegible] a[illegible] [illegible]atl. Acad. Sci. 27:4350-4354 (1979)] we[illegible] [illegible] [illegible]de preparations of GAG55 (30 ng [illegible] [illegible]ta[illegible] [illegible]protein per lane were typically used for [illegible] [illegible]t analysis. Following electrophoretic tra[illegible], [illegible]trocellulose strip were preincubated 30 [illegible] room temperature in Blotto buffer [illegible] [illegible]on-fat milk solids, 0.05% Tween 20, [illegible] [illegible] goat serum, phosphate buffered [illegible] [illegible]s were then incubated with a d[illegible] [illegible] [illegible]:50 in Blotto buffer) of sample [illegible] [illegible] for 1 hour to overnight at 4°C. [illegible] [illegible]ose strips were then washed exhaus[illegible] [illegible] [illegible]sphate buffered saline containing 0.[illegible] [illegible]-20. Human immunoglobulins were [illegible] [illegible]ally using biotinylated anti-human [illegible] [illegible]lin, and avidinD-horseradish pero[illegible] [illegible] Vector Laboratories, Burlingame CA[illegible] [illegible]tion of sera for competition exp[illegible] [illegible] incubation of 10 to 20 µl of ser[illegible] [illegible] [illegible]olumes of E. coli extract (containi[illegible] [illegible] bacterial protein) for 2 hours at 25°[illegible] [illegible] centrifugation to remove insoluble m[illegible].

Protein Purification

E. coli-expressed GAG55 was purified by immunoaffinity chromatography. Induced E. coli were collected by centrifugation and the cell pellet was stored at -70°C. The induced E. coli cell pellet was suspended in PBS and the cells were lysed using a French press. The lysate was centrifuged (27000 x g) and the supernatant applied to an immunoaffinity column containing $p17^{gag}$ mouse monoclonal-specific antibody BT2 (Biotech Research Laboratory, Rockville, MD) covalently attached to Reacti-gel (Pierce Chemical, Rockford, IL). The immunoaffinity column was prepared as specified by the supplier and washed and equilibrated with phosphate buffered saline (PBS). The E. coli cell lysate was loaded onto the column and the column was extensively washed with PBS. Fractions containing gag protein were eluted from the column using 0.5 N acetic acid and were then pH neutralized, lyophilized, and dissolved in PBS containing 0.1% SDS.

Alternatively, GAG55 was purified from a French press lysate of E. coli by cation exchange chromatography. Soluble proteins were dialyzed against 0.05 M phosphate buffer (pH 7.0), applied to a Pharmacia Mono S column (Pharmacia, Piscataway, NJ) and eluted with a NaCl gradient. The purity of the antigen was approximately 90% purified.

The peptide ENV9 was purified from E. coli containing the vector pENV9, as discussed in assignee's copending application, serial number 010056, filed Feb. 2, 1987. The predicted protein product of pENV9, designated ENV9, contains 340 amino acids which correspond to 54 amino acids from the N-terminal of the poliovirus sequence of pEXC [Ivanoff et. al., op. cit.], 46 amino acids of the C-terminal of the HIV env gp120 domain and 340 amino acids of the N-terminal of the env gp41 domain.

Enzyme Linked Immunosorbant Assay (ELISA)

Purified GAG55 antigen in 60 mM Na_2CO_3 pH 9.6 buffer containing 0.0005% SDS was applied to Immulon II microtiter plates at 100 ng/well at 4°C for about 18 hours. Purified ENV9 antigen was similarly applied in 60 mM carbonate pH 9.6 buffer containing 0.00006% SDS at a concentration of 20 ng/well. ELISA procedures were carried out as specified for the commercial HIV-ELISA kit (Du Pont, Wilmington, DE). The antigen-coated plates were washed with PBS + 0.05% Tween 20 (PBS-T). Plate washings were performed using 3 cycles on a Titertek Microplate Washer 120 followed by rotating the plate and washing again. The plates were incubated with PBS-T for 1 hour at 37° and were then washed 3 times with PBS-T and stored dry at 4° until they were used.

The plates were incubated with patient sera at

a 1:20 to 1:100 dilutio[illegible] PBS-T + 5% bovine serum albumi[illegible] inactivated normal goat serum (NGS[illegible] azide, 0.05% thimerosal) in the mic[illegible] incubated for 2 hours at 25°. The [illegible] washed with PBS-T, exposed to g[illegible] conjugated to alkaline phos[illegible] kson ImmunoResearch, West [illegible] hour at 25° and washed with PBS-[illegible] phosphatase reaction was develop[illegible] by the supplier, by exposure [illegible] nitrophenyl-phosphate in 100 µl [illegible] buffer (2 M, pH9.8) with magnesi[illegible] 0.02% sodium azide at a for 3 [illegible] followed by addition of sodium hy[illegible] centration of 1 N. The plates were [illegible] Multiskan MCC microtiter plate re[illegible] ngth of 405 nm.

BRIEF DESC[illegible]ES

FIGURE 1: Nuc[illegible] of the gag region of HIV_{BH10} (Ra[illegible] The start of the gag $p55^{gag}$ cod[illegible] indicated by an open bracket an[illegible] nd of the gag $p55^{gag}$ coding seque[illegible] by a closed bracket. The junctio[illegible] $p17^{gag}$ and $p24^{gag}$ polypeptides, t[illegible] olytic processing of $p55^{gag}$, is[illegible] $p55^{gag}$ is comprised by gag a[illegible] 1 to 512; $p17^{gag}$ is comprised b[illegible] number 1 to 132.

FIGURE 2: Sc[illegible] tion of E. coli plasmid vector de[illegible] full length, non-fused HIV gag pr[illegible] $p55^{gag}$. Plasmid pGAG55 (pgag5[illegible] s derived from pBR322 (thin lin[illegible] thick cros-shatched line), and [illegible] line), as well as a chemically [illegible] segment (thick closed line). In [illegible] on of the gag coding sequenc[illegible] the phage lambda P_L promoter [illegible] iation signals derived from a [illegible] ized DNA sequence. As indicat[illegible] es 5′ and preceding the gag A[illegible] n HIV_{BH10} have been changed i[illegible] eby generating a unique NcoI [illegible] ase site.

FIGURE 3: Ex[illegible] n E. coli. The accumulation of [illegible] was analyzed by SDS-PAGE [illegible] sates and Coomassie staining. L[illegible] 0 containing pBF106 (DC550/[illegible] nalidixic acid. Lane 2: DC55[illegible] ed. Lane 3: DC550/pGAG55, in[illegible] acid. The GAG55 protein represents approximately 5% of total cell protein in the nalidixic acid-induced DC550 cells containing pGAG55 (lane 3).

FIGURE 4: Immunoblot reactivity of E. coli-expressed GAG55-derived proteins with human AIDS patient serum antibodies. Total E. coli lysates were analyzed. Lane 1: DC550/pGAG55, non-induced; lane 2: DC550/pGAG55 nalidixic acid-induced; lane 3: DC550/pBF106, nalidixic acid-induced.

FIGURE 5: Binding of gag-specific antibodies in human HIV-seropositive sera by E. coli-expressed GAG55, as analyzed by competition binding and immunoblot. Prior to immunoblot of HIV antigens (Biotech Research Labs, Rockville, MD), human HIV-seropositive sera were preincubated with cell extracts prepared from induced E. coli, as indicated. Lane 9: DC550/pBF106, 20 µl extract; lane 10: DC550/pGAG55, 5 µl extract; lane 11: DC550/pGAG55, 20 µl extract. As shown, E. coli-expressed GAG55 protein effectively and selectively binds to all of the immunoblot-detectable gag-specific antibodies in this AIDS patient serum. Similar results have been obtained for several other HIV-seropositive sera. Thus, essentially all of the HIV gag-derived determinants and epitopes recognized by human antibodies are presented by E. coli-expressed GAG55. The figure also illustrates the identification of human serum samples containing HIV gag-reactive, HIV env- and pol-nonreactive antibodies. Human sera were analyzed by immunoblot of HIV antigens. In lanes 2 to 8, the specificity of the immunoblot reactivity for gag antigen was confirmed by competition with E. coli-expressed GAG55 protein (not shown).

FIGURE 6: Frequency distribution of GAG55-ELISA reactivity in three sample populations. Histograms show the fractions of samples in the absorbance ranges indicated on the horizontal axis. Absorbance ratio is ratio of the sample absorbance to that of a reference HIV positive control serum. Samples with absorbance ratios greater than 0.5 are grouped together for clarity. Upper panel: 69 normal HIV-ELISA-nonreactive blood donors. Middle panel: 57 representative HIV-ELISA reactive env-nonreactive sera from a blood donor population. Lower panel: 31 HIV positive sera: 26/31 of these sera are from patients with clinically diagnosed AIDS or AIDS related complex. Statistics for the populations were: normal sera, mean absorbance ratio = 0.00, s.d. = 0.03; HIV-reactive blood donors, mean = 0.13, s.d. = 0.11; HIV positives, mean = 0.37, s.d. = 0.43.

FIGURE 7: Comparison of apparent HIV gag-specific and env-specific antibody levels in human sera analyzed by ELISA. Antigens ENV9 and GAG55 were purified following expression in E. coli Symbols: open triangles, HIV-seronegative; open

circles, non-AIDS, HIV-serop ve; [illegible] uares, AIDS patients HIV-seroposit App [illegible] / gag-specific antibody levels are [illegible] c ced in AIDS patient sera compar [illegible] HIV-seropositive sera. This app [illegible] re ive [illegible] rease in gag-specific antibody le [illegible] since env-specific antibody level [illegible] ly affected.

EXAMP[illegible]

The invention is furthe [illegible] ri [illegible] ne following Examples, wherein [illegible] art [illegible] rcentages are by weight and deg [illegible] are Cel[illegible]

EXAMP[illegible]

Plasmid Constructions

HIV sequences were ob[illegible] d fr [illegible] roviral clone λ-BH10 (Ratner et [illegible] HIV-containing SacI fragment of [illegible] as isolated and treated with DNA[illegible] ic [illegible] ions of limited digestion to gener[illegible] and [illegible] fragments. The randomly cut H[illegible] A [illegible] s were blunt ended by Klenow rea[illegible] lig [illegible] HindIII linker, cut with HindIII, and i[illegible] ted [illegible] to unique HindIII site of plasmid pTC[illegible] P[illegible] TORF2 is derived from pBR322 ([illegible] 7); sequences inserted at the [illegible] te are transcribed under the contr[illegible] E. coli tryptophan operator and p[illegible] er et al., op. cit.). The resulting pla[illegible] were used to transform E. coli [illegible] re [illegible] containing HIV-derived seque[illegible] w [illegible] ed by colony hybridization using [illegible] HIV$_{BH10}$ proviral DNA as a probe [illegible] se [illegible] smids containing HIV-derived se[illegible] gnated pGAG3, pGAG4, PR2. pG[illegible] PR2 were shown by DNA sequ[illegible] contain HIV-derived DNA inser[illegible] g to nucleotides 227 to 1010, [illegible] 29 to 2037, respectively. HIV n[illegible] bered according to Ratner et al. [illegible] HIV$_{BH10}$ gag coding sequence [illegible] within nucleotides 334 to 1869 ([illegible] GAG3 and pGAG4 contain the 5 [illegible] oding sequence and the HIV-de[illegible] 2 extends beyond the 3′ end [illegible] g sequence.

Plasmids pGAG9 and [illegible] ecify the expression in E. coli un[illegible] moter control of the complete HIV$_{BH10}$ gag coding sequence, were derived as described below. pGAG9 was derived by ligation of the following DNA fragments: 1) the PstI to PstI tryptophan promoter-containing fragment of pGAG4, 2) the PstI to PstI pBR322 origin-containing fragment from pR2. pGAG10 was similarly derived by ligation of the following DNA fragments: 1) the Pst to PstI tryptophan promoter-containing fragment of pGAG3, 2) the PstI to PstI pBR322 origin-containing fragment of pR2. In this way the PstI site within the gag coding sequence was used to combine the 5′ of the gag-coding sequence from pGAG3 and pGAG4 with the 3′ of the gag coding sequence of pR2, to generate pGAG10 and pGAG9, respectively. pGAG10 places HIV sequence from nucleotides 227 to 2037 under transcriptional control of the tryptophan promoter. pGAG9 places HIV sequence from nucleotides 273 to 2037 under tryptophan promoter control (see Figure 1). E. coli translational control signals for translation initiation at the gag ATG at nucleotide 334 are provided by the HIV sequences immediately 5′ of the gag ATG. These translation initiation signals were predicted to be functional in E. coli on inspection of the HIV sequences.

Plasmid pGAG55 (Figure 2) is constructed by ligation of the following three DNA segments: 1) the large P$_L$-containing BamHI to NcoI fragment from pBF106; 2) the HIV gag-containing ClaI to BamHI DNA segment from pGAG9 or pGAG10; 3) a NcoI to ClaI oligonucleotide of the following sequence:

5′- C ATG GGT GCT AGA GCG TCA GTA TTA AGC GGG GGA GAA TTA GAT-3′ 3′- CCA CGC TCT CGC AGT CAT AAT TCG CCC CCT CTT AAT CTA GC-5′ The NcoI to ClaI oligonucleotide contains sequence information coding for the first 14 N-terminal amino acids of gag. Plasmid pGAG55 thereby contains the complete HIV$_{BH10}$ gag coding sequence fused to the synthetic E. coli translation initiation signal in plasmid pBF106, as described previusly (see p. 14). Plasmid pGAG55 is also referred to as pBF128. A unique NcoI site has been engineered at the translation start ATG of gag without altering the gag coding sequence (see Figure 2). The complete gag coding sequence can now be conveniently moved from GAG55 as a NcoI to BamHI DNA fragment cassette to other expression vectors. For example, the gag coding sequence can be precisely fused to signals directing the expression of the full length gag precursor in mammalian cells. Importantly, HIV sequences 5′ to the gag coding sequence have been eliminated in the NcoI to BamHI DNA cassette. HIV sequences immediately 5′ to the gag coding sequence, including a splice donor signal, are very likely to limit expression of gag in mammalian cells.

Plasmid GAG17, which specifies the expres-

sion in E. coli of the [illegible] ture, [illegible] d $p17^{gag}$ protein, was constructed [illegible] lig [illegible] following three DNA sequences: [illegible] rived, approximately 2962 base [illegible] DNA fragment from pGAG55 (see Figu [illegible] a PstI to PvuII fragment from pGAG55 ([illegible] tely 3068 base pair) containing pBF [illegible] ed DNA, lambda-derived DNA, a [illegible] he [illegible] f HIV gag (see Figure 2); 3) a che [illegible] lly sy [illegible] d PvuII to SalI DNA segment of the [illegible] nce (only 1 DNA strand is shown):

5′- CT GAC ACA GGA [illegible] AG GTC AGC CAA AAT TAC T [illegible] s [illegible] ull to SalI fragment extends the gag cod [illegible] nce from the PvuII site at nucleot [illegible] e 1) and introduces of TAA stop [illegible] don [illegible] yr codon (TAC) at the $p17^{gag}$-p2 [illegible] ur [illegible] e 1). The 3′ G residue of the Indica [illegible] s the first nucleotide of a SalI site In [illegible] a coding sequence correspondin [illegible] cleotides 334 to 729 of HIV_{BH10} [illegible] ced under transcriptional and translation [illegible] derived from pBF106. pBF106 is [illegible] re [illegible] usly.

EXAM[illegible]E [illegible]

Reactivity of E. coli-e[illegible] with HIV-Specific Antibodies

As shown in Figure [illegible] n can be efficiently expressed a [illegible] coli, using the appropriate plasmid [illegible] re [illegible] system. Analysis of total E. coli [illegible] eins [illegible] AGE and Coomassie blue staini [illegible] AG55 accumulates to a level [illegible] of total E. coli protein. This [illegible] ssion of GAG55 facilitates pu [illegible] atio [illegible] protein. GAG17 similarly accu[illegible] vel in E. coli.

Monoclonal antib[illegible] HIV-expressed $p17^{gag}$ (BT2, [illegible] ch [illegible] abs) and $p24^{gag}$ (BT3, Biotech R[illegible] rch [illegible] found to react specifically with [illegible] 5 [illegible] en tested in immunoblot and E[illegible] methods described previously. [illegible] bbit antiserum specific for h[illegible] $p24^{gag}$ - (obtained from D. R[illegible] erimental Station) was immunore[illegible] xpressed GAG55. Furthermore, [illegible] GAG55 was shown to be im[illegible] many human AIDS patient ser[illegible] iot assay (for example, see Fig[illegible] coli-expressed GAG17 is im[illegible] $p17^{gag}$-specific monoclonal antibody BT2 and many human AIDS patient sera. These results confirm that viral antigenic sites reside within E. coli-expressed GAG55 and GAG17.

EXAMPLE 3

Use of E. coli-expressed GAG55 as an Immunogen

Another advantage of the antigenic structure of E. coli-expressed GAG55 is the ability of the peptide to illicit HIV-specific antibodies in animals. Purified GAG55 was injected into rabbits and goats and the resulting antisera were assayed by immunoblot and ELISA, according to the methods described previously. Goat and rabbit antisera to E. coli-expressed GAG55 reacted with the expected HIV-expressed gag-derived proteins $p55^{gag}$, $p24^{gag}$, and $p17^{gag}$ on viral immunoblot. Antisera to GAG55 also reacted with an HIV-ELISA (Du Pont).

The titer of the antisera to E. coli-expressed GAG55 was found to be equal to or greater than antisera to gag proteins purified from HIV in specifically binding to the HIV-expressed gag proteins. This result demonstrates that E. coli-expressed GAG55 contains the major native viral epitopes and that antibodies to GAG55 are sensitive reagents for detecting viral proteins.

EXAMPLE 4

Detection of Antibodies to HIV gag in Human Sera

The E. coli-expressed gag-derived protein GAG55 is comprised of protein sequences covering the complete gag open reading frame and should, therefore, be very similar antigenicaly to $P55^{gag}$ expressed by HIV in infected mammalian cells. In order to compare the ability of GAG55 and HIV-expressed gag antigens to be recognized by human antibodies in HIV-seropositive sera, experiments involving competition between recombinant and viral gag antigens for reaction with serum antibodies were carried out. Using a competition immunoblot technique, six HIV-seropositive human sera were tested for reactivity with GAG55 antigen as described below. Sera were pre-absorbed with GAG55-containing E. coli extract and, in parallel, with a control E. coli extract. These absorptions were followed by analysis of the sera using HIV

antigen strips (Biotech Rese[illegible] Labs [illegible] each HIV positive serum tested, [illegible]mbina[illegible] gag poly-protein was found to substan[illegible]y, if not co[illegible]plete-ly, out-compete reactivity [illegible]ll the gag-[illegible]iated immunoblot bands (see Figu[illegible] 5). The spec[illegible]ity of this competition effect for gag was sh[illegible] the lack of competition of env or pol bands [illegible] the lack of effect of control E. c[illegible]xtracts [illegible] HIV immunoblot band. The pres[illegible]sults [illegible]strate that most antibodies in HIV-[illegible]positi[illegible] ser[illegible] reac-ting with HIV gag immunobl[illegible]nds also re[illegible]t with GAG55 antigen. Thus, E. co[illegible]pressed G[illegible]55 is clearly a useful reagent fo[illegible] detection [illegible] HIV gag-specific antibodies and, therefore, [illegible] be a useful component of HIV [illegible]nostic ar[illegible]od screening procedures. The[illegible]cedu[illegible] in-volve, for example, immuno[illegible] ELIS[illegible]m[illegible]ods.

EXAMP[illegible]

Purification of gag-specific [illegible]ies in [illegible]man HIV-seropositive serum by [illegible]y c[illegible]ato[illegible]raphy using E. coli-expressed GAG[illegible]

E. coli-expressed GAG[illegible] was demo[illegible]ated using immunoblot assays t[illegible] to [illegible] all of the gag-specific antibodi[illegible] many [illegible]IV-seropositive sera (Example [illegible]ure 5 [illegible] T[illegible]ore, we prepared an affinity col[illegible] the pr[illegible]tive isolation of HIV gag-spec[illegible]s. [illegible]355 was partially purified as [illegible]ribe[illegible] and covalently coupled to a Re[illegible]el support [illegible]ierce Chemical, Rockford, IL) as [illegible]ified by t[illegible]up-plier. Human serum was lo[illegible] to a [illegible] con-taining the GAG55 affinity [illegible] was washed extensively using [illegible] human antibodies were el[illegible] the eluted fractions were pH [illegible] lot analysis of HIV antigens ([illegible] L[illegible]bs, Rockville, MD) using the [illegible]55-puri[illegible]ied [illegible]uman antibodies revealed selecti[illegible] immunoreac[illegible] of these antibodies with gag-[illegible] ese results demonstrate the uti[illegible] E. co[illegible]ssed gag protein for the prep[illegible] gag-specific antibodies.

HIV gag-specific antib[illegible] re-search reagents for the id[illegible]ng of gag-specific epitopes an[illegible] host immune response to HIV. [illegible] it [illegible] be possible to identify gag-sp[illegible] can neutralize or block HIV inf[illegible] of such HIV- neutralizing an[illegible]tic reagents will depend on [illegible] for their purification. As disc[illegible] full length gag protein, such as E. coli-expressed GAG55, that represents essentially all of the gag-specific epitopes may be the preferred affinity reagent for the identification and purification of gag-specific antibodies. Subfragments of gag protein, as previously disclosed by others, can only represent a subset of the total set of gag-specific epitopes.

EXAMPLE 6

Identification of HIV gag-reactive, env- and pol-nonreactive Sera

ELISAs utilizing as antigen purified HIV grown in tissue culture (HIV-ELISA) have proved valuable for screening blood for evidence of exposure to HIV. Many times sera will repeatedly test positive on HIV-ELISA but will not show clear evidence of HIV exposure when subsequently examined by HIV-immunoblot. During screening of blood donors, we and others [Courouce et al., Lancet II:921-922 (1986); Biberfeld et al., Lancet II:289-298 (1986)] often detect reactivity in HIV immunoblots at positions of gag antigens p15, p17, p24, or p55 unaccompanied by detectable immunoreactivity with env. Examples of such HIV gag-reactive, env-nonreactive sera are shown in Figure 5. The specificity of these human antibodies for HIV gag was demonstrated by competition binding with GAG55. Thus, preincubation of the sera with GAG55 specifically blocked binding of the antibody to the gag proteins on the HIV-immunoblot.

More evidence demonstrating the occurence of anti-gag antibodies and quantitative data on their relative levels in different populations was obtained using purified recombinant antigens in ELISA format as shown in Figure 6. The HIV-ELISA-reactive blood donors are distinct from the normal group in having elevated gag ELISA titres, as does the HIV-positive group. In addition, the GAG55-ELISA values for the HIV-ELISA-reactive blood donors are in the same range as several of the HIV-positive sera. Analysis of these samples with the ENV9-ELISA revealed that the absorbances of the HIV-ELISA reactive donor group were distinctly lower than the HIV-positive group and similar to those of normal sera.

For a high percentage of the more than 200 HIV-ELISA positive sera that we have examined, there is unequivocal evidence of antibodies that are immunoreactive with HIV gag antigens, unaccompanied by evidence of HIV env reactivity. The methods demonstrated utilizing both GAG55-ELISA

and ENV9-ELISA sho[illegible] [illegible] routine blood screening and diagnostic a[illegible].

The ability to identify such [illegible]g-seropositive and env-seronegative human [illegible] samples is of considerable impor[illegible] several studies, including that [illegible] presented at the National Institute[illegible] [illegible]us Development Conference [illegible] Routine HTLV-III Antibody Tes[illegible]" (July 7-9, 1986), have indicate[illegible] [illegible]ted individuals who initially display [illegible] reactivity only to gag specific b[illegible] [illegible]quently display definite HIV ser[illegible] [illegible]nced by env reactivity. It is also [illegible] [illegible]V gag-reactive, env-nonreactive se[illegible] [illegible]ure to other human retrovirus[illegible] [illegible]TLV-II or HIV-II, and possibly [illegible] retroviruses. The significan[illegible] [illegible]ransfusion recipient of such HIV gag[illegible] env-nonreactive sera remains to be [illegible] and further studies are essential. [illegible] blood from donors with thes[illegible] should not be transfused. This [illegible] [illegible]oth the importance of methods [illegible] [illegible]V gag proteins and the utility [illegible] [illegible]ose.

EX[illegible]

Analysis of Human Ant[illegible] [illegible] coli-Expressed env- and ga[illegible]

Although the HIV [illegible] [illegible]werful and sensitive tool for scre[illegible] for the presence of HIV-speci[illegible] [illegible]y utilizing total virus antige[illegible] [illegible]ormation concerning which [illegible] [illegible]nents of the virus are immu[illegible] sera. Immunoblot or immu[illegible] [illegible]s can be used to identify [illegible] [illegible]y to [illegible]pecific viral components. To [illegible] [illegible]ition [illegible] HIV-seropositivity is base[illegible] [illegible]ot reactivity with both gag- and env[illegible].

As an alternative [illegible] im-munoprecipitation met[illegible] [illegible] relative levels of antibodi[illegible] [illegible] and env proteins, we hav[illegible] [illegible] env-specific ELISAs, usin[illegible] [illegible]sed GAG55 and ENV9 as antigens. [illegible] amino acids corresponding to the [illegible] [illegible]s of [illegible]e C-terminus of the $HIV_{BH1[illegible]}$ [illegible] p1[illegible] [illegible]ain [illegible]d 240 amino acids of the N-[illegible] [illegible] main. Methods for detection [illegible] relative levels of antibodi[illegible] [illegible] antigenic components o[illegible] [illegible] env-encoded proteins, provides information that may be useful for the diagnosis and prognosis of HIV-infected individuals. The ELISA format provides considerable advantages compared to immunoblot and immunoprecipitation methods. The ELISA procedure is sensitive, quantitative, rapid, non-radioactive, and relatively inexpensive and simple to run.

Recently several ELISA assays have been developed which utilize as antigen proteins expressed in E. coli to detect antibodies to the AIDS virus [Dowbenko, et al., Proc. Natl. Acad. Sci. 82: 7748-7752 (1985); Steimer, et al., Virology 150:283-290 (1986); Cabradilla et al., Biotechnology 4: 128-133 (1986); Chang et al., Biotechnology 3:905-909 (1985); Shoeman et al., Anal. Biochem. 161:370-379 (1987)]. Previous recombinant gag-derived ELISAs which have been described only contain subregions of gag, such as $p24^{gag}$, and do not contain sequences corresponding to the entire gag coding sequence. Since the GAG55 represents the complete gag coding sequence, the GAG55-ELISA may detect HIV gag-specific antibodies in human sera that are not detected by ELISAs employing subregions of gag. In particular, GAG55 is preferred to other reported recombinant gag antigens because the complete $p17^{gag}$ region is represented. In addition, GAG55 and GAG17 do not contain any non-HIV or non-HIV gag amino acid sequences. These antigens are preferred for antibody detection because of the specificity for gag and of the reduced probability of detecting false-positive, non-HIV immunoreactivity.

The HIV gag-specific and env-specific antibody levels were quantitated in human sera from individuals of known clinical status, using both the GAG55-ELISA and ENV9-ELISA. Sera categorized as being from AIDS patients, HIV-seropositive (non-AIDS), and HIV-seronegative individuals were obtained from Duke University. These sera were analyzed using the GAG55 and ENV9 ELISAs. The results of these assays are shown in Figure 7.

As shown in Figure 7, the gag-specific apparent antibody levels in many of the AIDS patient sera are reduced compared to the gag-specific apparent antibody levels in non-AIDS, HIV-seropositive sera. This effect is selective, since the HIV env-specific antibody levels do not distinguish the AIDS and non-AIDS seropositive groups. The GAG55-ELISA alone is obviously limited as an HIV screening assay since many of the characterized HIV positive sera were not detected as positive (see also Example 6, Figure 6). This does not, however, preclude the use of a gag-specific ELISA as an important component of diagnostic assays. For example, many sera exhibit HIV gag immunoreactivity without detectable env-specific immunoreactivity (Example 6). Moreover, quantitation of the relative antibody levels specific to gag and

env and other HIV antigenic components may be useful for monitoring HIV-associated disease. The results presented indicate that it may be possible to develop a correlation between the serum reactivity on the GAG55 and ENV ELISA assays and the stage of disease. The present data provide evidence that quantitation by ELISA of the relative HIV env-specific and gag-specific antibody levels in human sera is useful for both screening and diagnostic analysis of human sera.

Claims

1. A recombinant plasmid capable of high level expression in E. coli of a nonfused peptide corresponding to the complete HIV p55 coding sequence and exhibiting the antigenicity of the complete HIV $p55^{gag}$ protein, consisting essentially of plasmid pBR322 and DNA coding for amino acids 1 to 512 of HIV protein $p55^{gag}$, with a translation start signal and an ATG/Met translation start codon downstream from a promoter selected from the group consisting of the E. coli lac operon promoter, the E. coli tryptophan promoter, and the phage lambda P_L promoter.

2. A recombinant plasmid of Claim 1 having incorporated therein a translation start signal which is an HIV sequence immediately 5' of the gag ATG/Met translation start codon or an oligonucleotide linker as described in Example 1.

3. A recombinant plasmid of Claim 1 wherein the gag ATG/Met translation start codon is part of an NcoI restriction endonuclease site.

4. An E. coli cell transformed with a plasmid of anyone of claims 1 to 3.

5. An HIV gag-encoded peptide expressed by a cell of Claim 4.

6. A recombinant plasmid capable of high level expression in E. coli of a nonfused peptide corresponding to the complete p17 subregion of the HIV $p55^{gag}$ coding sequence and exhibiting the antigenicity of the complete HIV $p17^{gag}$ protein, consisting essentially of plasmid pBR322 and DNA coding for amino acids 1 to 132 of HIV protein $p55^{gag}$, with a translation start signal and an ATG/Met translation start codon downstream from a promoter selected from the group consisting of the E. coli lac operon promoter, the E. coli tryptophan promoter, and the phage lambda P_L promoter.

7. A recombinant plasmid of Claim 6 having incorporated therein a translation start signal which is an HIV sequence immediately 5' of the gag ATG/Met translation start codon or an oligonucleotide linker as described in Example 1.

8. A recombinant plasmid of Claim 6 wherein the gag ATG/Met translation start codon is part of an NcoI restriction endonuclease site.

9. An E. coli cell transformed with a plasmid of anyone of claims 6 to 8.

10. An HIV gag-encoded peptide expressed by a cell of Claim 9.

11. An E. coli-expressed recombinant nonfused peptide corresponding to the complete HIV $p55^{gag}$ coding sequence and exhibiting the antigenicity of the complete HIV $p55^{gag}$ protein.

12. An E. coli-expressed recombinant nonfused peptide according to Claim 11 which is encoded by nucleotides numbered 334 to 1869 of the HIV genome, as shown in Figure 1.

13. An E. coli-expressed recombinant nonfused peptide corresponding to the complete p17 subregion of the HIV $p55^{gag}$ coding sequence, and exhibiting the antigenicity of the complete HIV $p17^{gag}$ protein.

14. An E. coli-expressed recombinant nonfused peptide according to Claim 13 which is encoded by nucleotides numbered 334 to 729 of the HIV genome, as shown in Figure 1.

15. In a diagnostic kit used for detecting antibodies to HIV in a biological sample wherein the sample is contacted with a peptide and immunoreactivity is detected, the improvement comprising employing a peptide of anyone of claims 5, 10, 11, and 13.

16. In a process for detecting antibodies to HIV in a biological sample, comprising contacting said sample with a peptide which is immunoreactive with said antibodies and detecting immunoreactivity, the improvement comprising employing a peptide of anyone of claims 5, 10, 11, and 13.

17. A vaccine protective against HIV comprising an effective protective amount of a peptide of anyone of claims 5, 10, 11, and 13 in a pharmaceutically acceptable carrier.

FIG. 1A

228 255
GAA AGC GGA A [illegible] AGC TCT CTC GAC GCA GGA CTC GGC TTG CTG AAG
Glu Ser Gly L[illegible] Ser Ser Leu Asp Ala Gly Leu Gly Leu Leu Lys

282 309
CGC GCA CGG [illegible] GG GGC GGC GAC TGG TGA GTA CGC CAA AAA TTT TGA
Arg Ala Arg Gln [illegible] Gly Gly Asp Trp . Val Arg Gln Lys Phe .

336 363
CTA GCG GAG [illegible] AG GAG ATG GGT GCG AGA GCG TCA GTA TTA AGC GGG
Leu Ala Glu [illegible] Ar[illegible] Glu MET Gly Ala Arg Ala Ser Val Leu Ser Gly

390 417
GGA GAA TTA [illegible] AAA ATT CGG TTA AGG CCA GGG GGA AAG AAA AAA
Gly Glu Leu [illegible] Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys

444 471
TAT AAA TTA [illegible] GT. TGG GCA AGC AGG GAG CTA GAA CGA TTC GCA GTT
Tyr Lys Leu [illegible] Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Val

498 525
AAT CCT GGC [illegible] TCA GAA GGC TGT AGA CAA ATA CTG GGA CAG CTA
Asn Pro Gly [illegible] Ser Glu Gly Cys Arg Gln Ile Leu Gly Gln Leu

552 579
CAA CCA TCC [illegible] GG TCA GAA GAA CTT AGA TCA TTA TAT AAT ACA GTA
Gln Pro Ser [illegible] Ser Glu Glu Leu Arg Ser Leu Tyr Asn Thr Val

606 633
GCA ACC CTC [illegible] CAA AGG ATA GAG ATA AAA GAC ACC AAG GAA GCT
Ala Thr Leu [illegible] Gln Arg Ile Glu Ile Lys Asp Thr Lys Glu Ala

660 687
TTA GAC AAG [illegible] CAA AAC AAA AGT AAG AAA AAA GCA CAG CAA GCA
Leu Asp Lys [illegible] Gln Asn Lys Ser Lys Lys Lys Ala Gln Gln Ala

714 741
GCA GCT GAC [illegible] AGT CAG GTC AGC CAA AAT TAC CCT ATA GTG CAG
Ala Ala Asp [illegible] Ser Gln Val Ser Gln Asn Tyr Pro Ile Val Gln
17 24

768 795
AAC ATC CAG [illegible] CAT CAG GCC ATA TCA CCT AGA ACT TTA AAT GCA
Asn Ile Gln [illegible] His Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala

822 849
TGG GTA AAA [illegible] AAG GCT TTC AGC CCA GAA GTA ATA CCC ATG TTT
Trp Val Lys [illegible] Lys Ala Phe Ser Pro Glu Val Ile Pro MET Phe

FIG. 1B

```
                                 876                                903
TCA GCA TTA TCA GAA GGA [illegible] ACC CCA CAA GAT TTA AAC ACC ATG CTA AAC ACA
Ser Ala Leu Ser Glu Gly [illegible] Thr Pro Gln Asp Leu Asn Thr MET Leu Asn Thr

                                 930                                957
GTG GGG GGA CAT [illegible] GCA G[illegible]C ATG CAA ATG TTA AAA GAG ACC ATC AAT GAG GAA
Val Gly Gly His [illegible] Ala Ala MET Gln MET Leu Lys Glu Thr Ile Asn Glu Glu

                                 984                               1011
GCT GCA GAA TGG GAT AGA GTA CAT CCA GTG CAT GCA GGG CCT ATT GCA CCA GGC
Ala Ala Glu Trp Asp Arg Val His Pro Val His Ala Gly Pro Ile Ala Pro Gly

                                1038                               1065
CAG ATG AGA GAA C[illegible] AGG G[illegible]A AGT GAC ATA GCA GGA ACT ACT AGT ACC CTT CAG
Gln MET Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr Ser Thr Leu Gln

                                1092                               1119
GAA CAA ATA GGA TGG ATG ACA AAT AAT CCA CCT ATC CCA GTA GGA GAA ATT TAT
Glu Gln Ile Gly Trp MET Thr Asn Asn Pro Pro Ile Pro Val Gly Glu Ile Tyr

                                1146                               1173
AAA AGA TGG ATA [illegible] CTG GGA TTA AAT AAA ATA GTA AGA ATG TAT AGC CCT ACC
Lys Arg Trp Ile [illegible] Leu Gly Leu Asn Lys Ile Val Arg MET Tyr Ser Pro Thr

                                1200                               1227
AGC ATT CTG GAC ATA AGA CAA GGA CCA AAA GAA CCT TTT AGA GAC TAT GTA GAC
Ser Ile Leu Asp Ile Arg Gln Gly Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp

                                1254                               1281
CGG TTC TAT AAA [illegible] CTA AGA GCC GAG CAA GCT TCA CAG GAG GTA AAA AAT TGG
Arg Phe Tyr Lys [illegible] Leu Arg Ala Glu Gln Ala Ser Gln Glu Val Lys Asn Trp

                                1308                               1335
ATG ACA GAA A[illegible] [illegible] TTG GTC CAA AAT GCG AAC CCA GAT TGT AAG ACT ATT TTA
MET Thr Glu Thr Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu

                                1362                               1389
AAA GCA TTG G[illegible] [illegible] GCG GCT ACA CTA GAA GAA ATG ATG ACA GCA TGT CAG GGA
Lys Ala Leu Gly Pro Ala Ala Thr Leu Glu Glu MET MET Thr Ala Cys Gln Gly

                                1416                               1443
GTA GGA GGA C[illegible] G[illegible] CAT A[illegible]G GCA AGA GTT TTG GCT GAA GCA ATG AGC CAA GTA
Val Gly Gly Pro Gly His Lys Ala Arg Val Leu Ala Glu Ala MET Ser Gln Val

                                1470                               1497
ACA AAT ACA G[illegible] [illegible] ATA ATG ATG CAG AGA GGC AAT TTT AGG AAC CAA AGA AAG
Thr Asn Thr Ala [illegible] Ile MET MET Gln Arg Gly Asn Phe Arg Asn Gln Arg Lys

                                1524                               1551
ATG GTT AAG T[illegible] T[illegible] AAT TGT GGC AAA GAA GGG CAC ACA GCC AGA AAT TGC AGG
MET Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His Thr Ala Arg Asn Cys Arg
```

F I G. I C

```
                                      1578                                      1605
GCC CCT AGG AAA AAG GGC TGT TGG AAA TGT GGA AAG GAA GGA CAC CAA ATG AAA
Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly Lys Glu Gly His Gln MET Lys

                                      1632                                      1659
GAT TGT ACT GAG AGA CAG GCT AAT TTT TTA GGG AAG ATC TGG CCT TCC TAC AAG
Asp Cys Thr Glu Arg Gln Ala Asn Phe Leu Gly Lys Ile Trp Pro Ser Tyr Lys

                                      1686                                      1713
GGA AGG CCA GGG AAT TTT CTT CAG AGC AGA CCA GAG CCA ACA GCC CCA CCA TTT
Gly Arg Pro Gly Asn Phe Leu Gln Ser Arg Pro Glu Pro Thr Ala Pro Pro Phe

                                      1740                                      1767
CTT CAG AGC AGA CCA GAG CCA ACA GCC CCA CCA GAA GAG AGC TTC AGG TCT GGG
Leu Gln Ser Arg [illegible] Glu Pro Thr Ala Pro Pro Glu Glu Ser Phe Arg Ser Gly

                                      1794                                      1821
GTA GAG ACA ACA [illegible] CCC CCT CAG AAG CAG GAG CCG ATA GAC AAG GAA CTG TAT
Val Glu Thr Thr [illegible] Pro Pro Gln Lys Gln Glu Pro Ile Asp Lys Glu Leu Tyr

                                      1848                                      1875
CCT TTA ACT TCC [illegible] AGA TCA CTC TTT GGC AAC GAC CCC TCG TCA CAA TAA AGA
Pro Leu Thr Ser [illegible] Arg Ser Leu Phe Gly Asn Asp Pro Ser Ser Gln  .  Arg

                                      1902                                      1929
TAG GGG GGC AAC [illegible] AGG AAG CTC TAT TAG ATA CAG GAG CAG ATG ATA CAG TAT
 .  Gly Gly Asn  .  Arg Lys Leu Tyr  .  Ile Gln Glu Gln MET Ile Gln Tyr

                                      1956                                      1983
TAG AAG AAA TGA [illegible] TGC CAG GAA GAT GGA AAC CAA AAA TGA TAG GGG GAA TTG
 .  Lys Lys  .  [illegible] Lys Gln Glu Asp Gly Asn Gln Lys  .   .  Gly Glu Leu

                                      2010
GAG GTT TTA TCA [illegible] TA
Glu Val Leu Ser Lys
```

FIG. 2

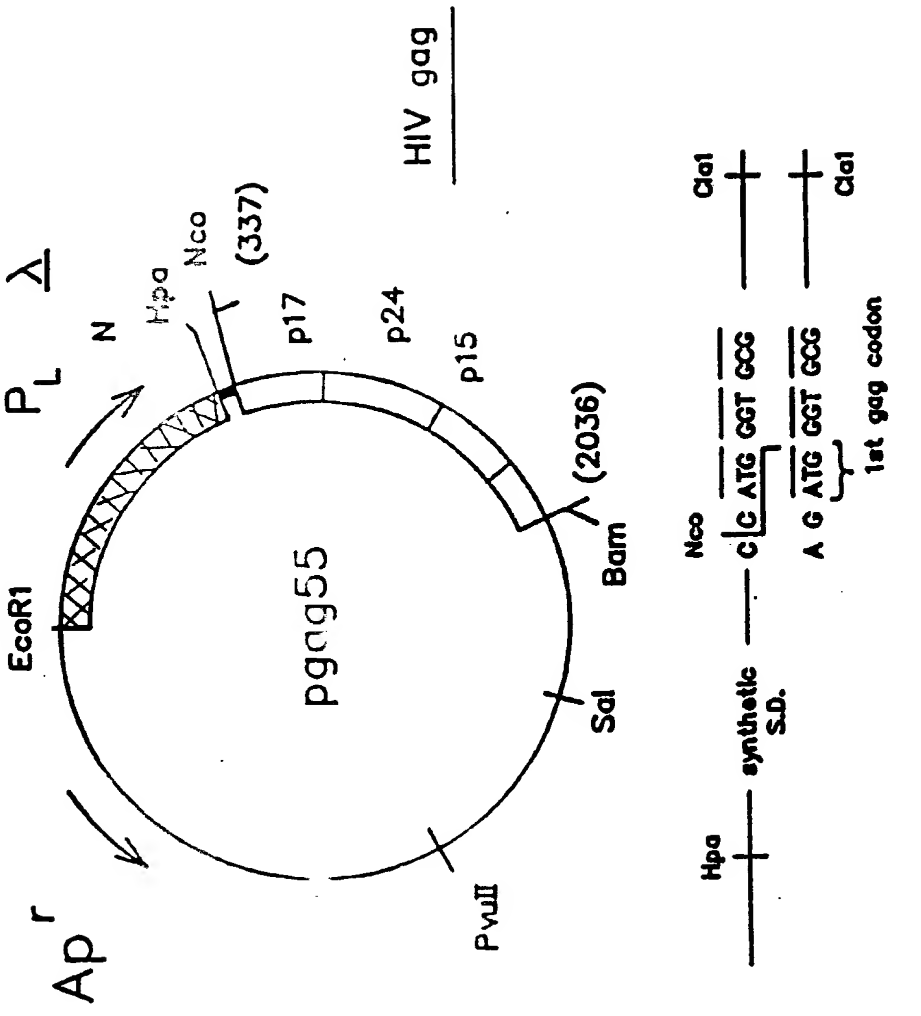

Ap r
P L
λ
N
Hpa
Nco
(337)
p17
p24
p15
(2036)
Bam
Sal
PvuII
EcoR1
pgag55
HIV gag
Hpa
synthetic S.D.
Nco
C C ATG GGT GCG
A G ATG GGT GCG
1st gag codon
Cla1
Cla1

FIG. 3

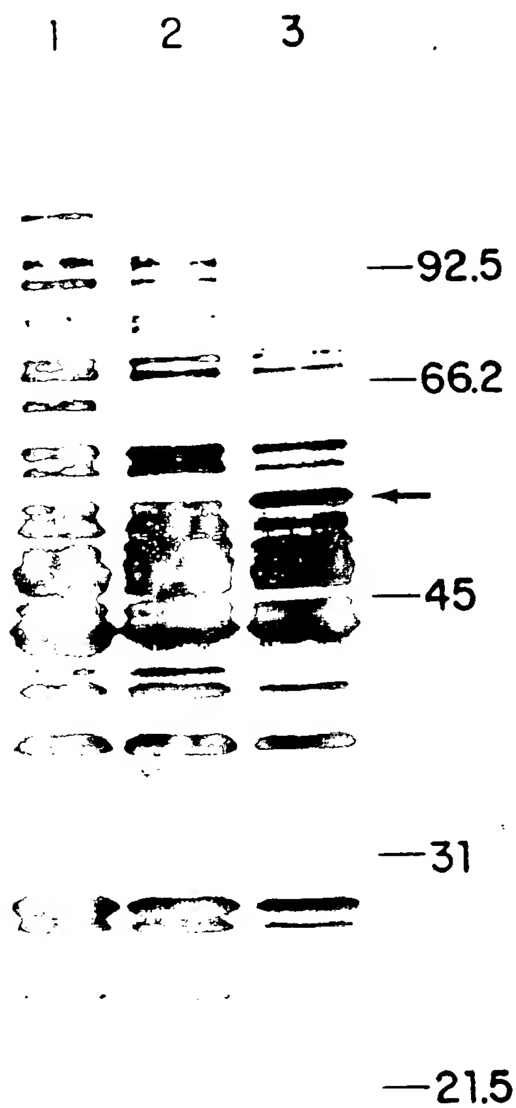
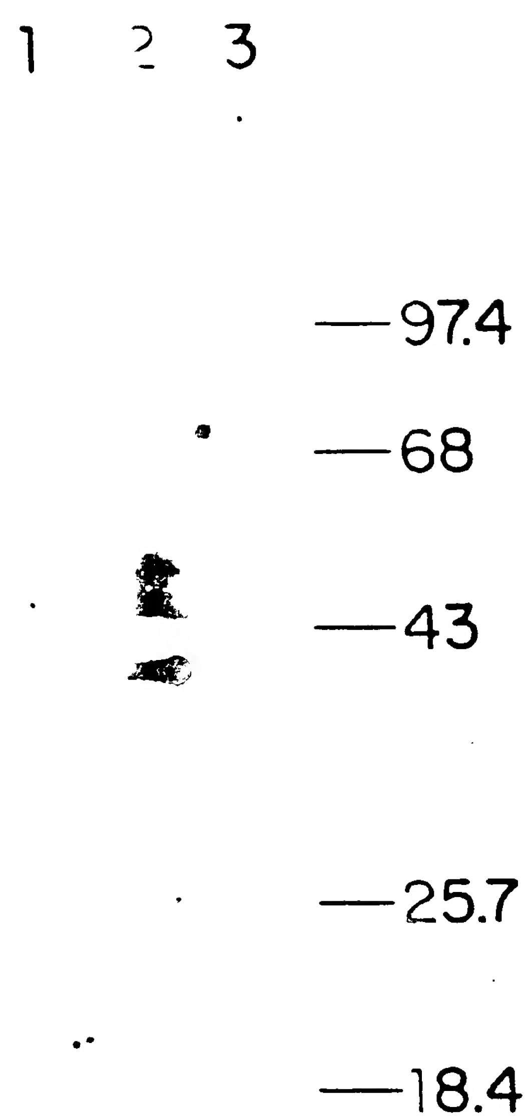


FIG. 4



F I G. 5

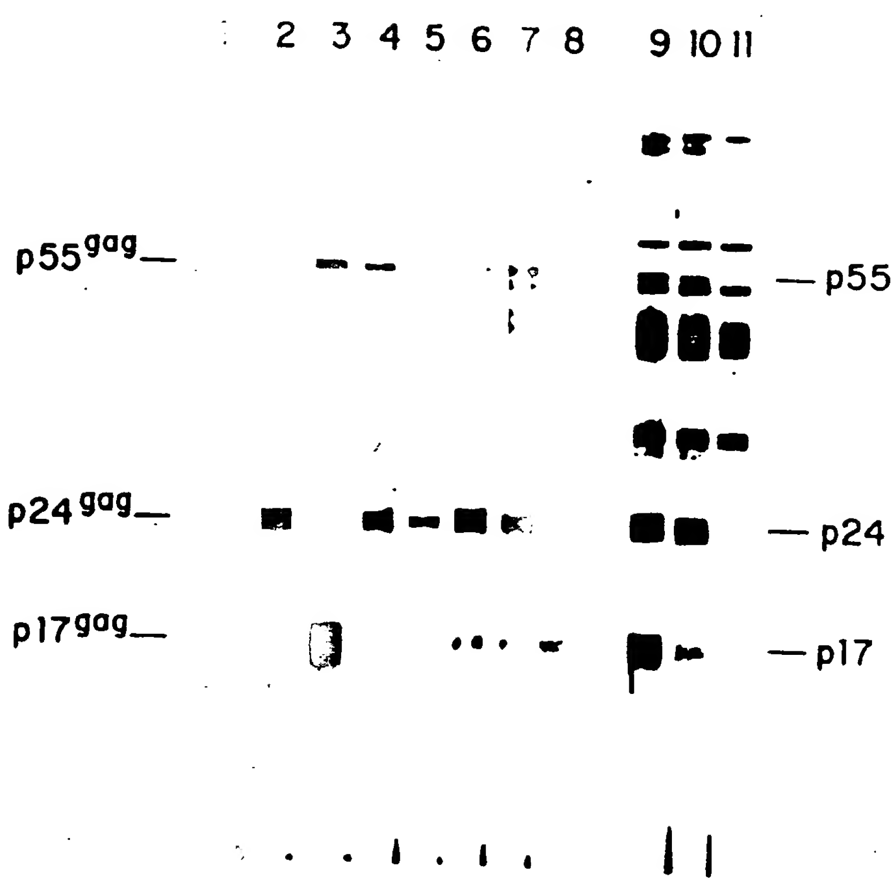


FIG. 6A

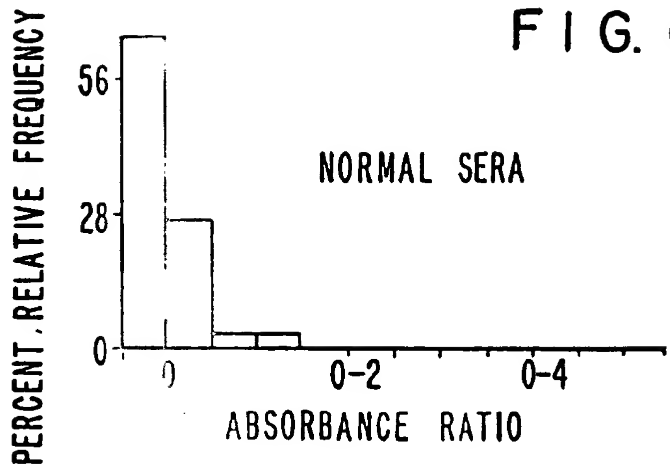


FIG. 6B

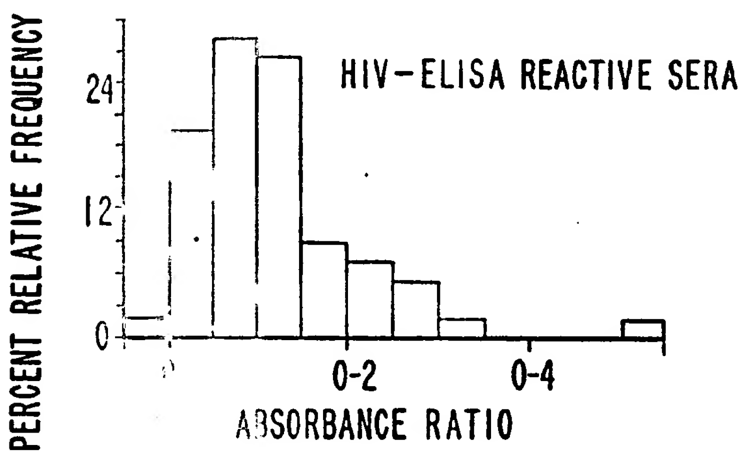
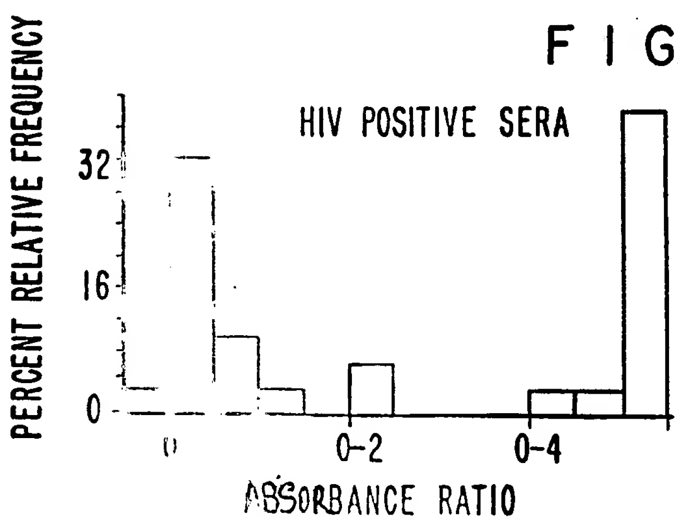
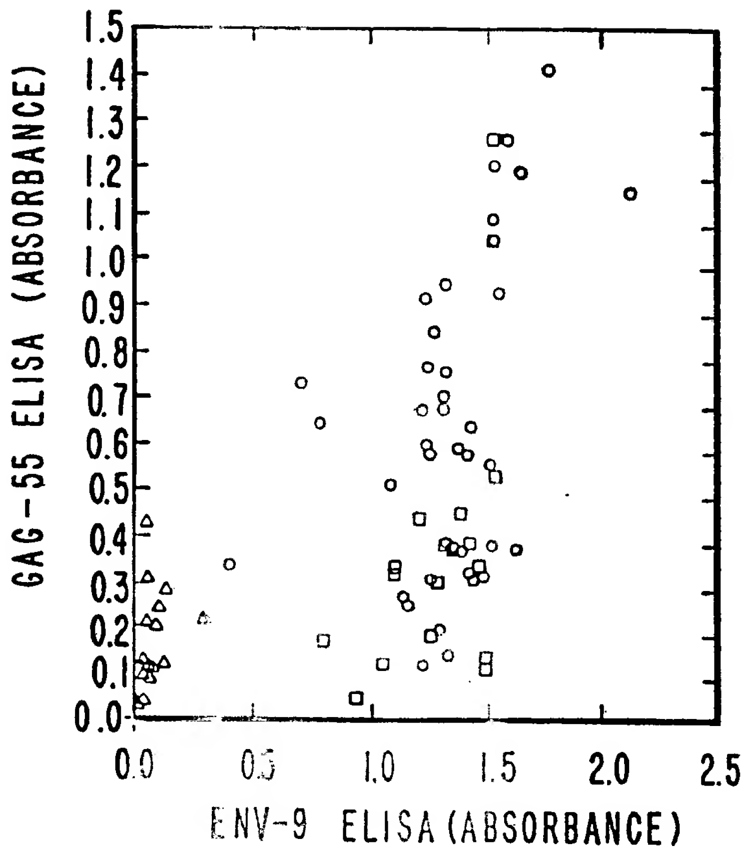


FIG. 6C



F I G. 7


GAG-55 ELISA (ABSORBANCE)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
1.1
1.2
1.3
1.4
1.5
0.0
0.5
1.0
1.5
2.0
2.5
ENV-9 ELISA (ABSORBANCE)